


```

FT DISULFID 385 418 BY SIMILARITY.
FT CARBOHYD 88 88 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 136 136 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 141 141 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 156 156 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 160 160 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 186 186 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 197 197 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 230 230 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 234 234 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 241 241 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 262 262 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 276 276 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 289 289 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 295 295 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 301 301 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 332 332 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 339 339 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 356 356 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 386 386 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 392 392 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 397 397 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 406 406 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 448 448 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 463 463 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 611 611 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 616 616 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 624 624 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 637 637 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 674 674 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 750 750 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 816 816 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 856 AA; 97212 MW; 6FAD16AF85107FE0 CRC64;

```

```

Query Match 75.7%; Score 112; DB 1; Length 856;
Best Local Similarity 85.7%; Pred. No. 3.3e-08;
Matches 24; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

```

OY 1 RVIRVQACRAIRHVRIRROGLRIL 28
Db 828 RVLEVQACRAIRHVRIRROGLRIL 855

```

```

RESULT 2
ENV_HVILW STANDARD; PRT; 856 AA.
AC 070626;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope glycoprotein gp160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (HIV-1).
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCB1_TaxID=82834;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95127297; PubMed=7826699;
RA Reitz M.S. Jr., Hall L., Robert-Guroff M., Lautenberger J., Hahn B.M.,
Shaw G.M., Kong L.I., Weiss S.H., Waters D., Gallo R.C., Blattner W.;
RT "Viral variability and serum antibody response in a laboratory worker
infected with HIV type 1 (HIV type IIB).";
RL AIDS Res. Hum. Retroviruses 10:1143-1155(1994).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

CC EMBL: U12055; AA76690.1; -
DR GMB: U12055; AA76690.1; -
DR InterPro: IPR000328; Env GP41.
DR InterPro: IPR000777; GP120.
DR Pfam: PF00516; GP120; 1.
DR Pfam: PF00517; GP41; 1.
KW AIDS; Coat protein; Glycoprotein; Transmembrane;
KW signal.
FT SIGNAL 1 30
FT CHAIN 31 511
FT CHAIN 512 856
FT DISULFID 54 74
FT DISULFID 119 205
FT DISULFID 126 196
FT DISULFID 131 157
FT DISULFID 218 247
FT DISULFID 228 239
FT DISULFID 296 331
FT DISULFID 378 445
FT DISULFID 385 418
FT CARBOHYD 88 88
FT CARBOHYD 136 136
FT CARBOHYD 141 141
FT CARBOHYD 156 156
FT CARBOHYD 160 160
FT CARBOHYD 186 186
FT CARBOHYD 197 197
FT CARBOHYD 230 230
FT CARBOHYD 234 234
FT CARBOHYD 241 241
FT CARBOHYD 262 262
FT CARBOHYD 276 276
FT CARBOHYD 289 289
FT CARBOHYD 295 295
FT CARBOHYD 301 301
FT CARBOHYD 332 332
FT CARBOHYD 339 339
FT CARBOHYD 356 356
FT CARBOHYD 386 386
FT CARBOHYD 392 392
FT CARBOHYD 397 397
FT CARBOHYD 406 406
FT CARBOHYD 448 448
FT CARBOHYD 463 463
FT CARBOHYD 611 611
FT CARBOHYD 616 616
FT CARBOHYD 624 624
FT CARBOHYD 637 637
FT CARBOHYD 674 674
FT CARBOHYD 750 750
FT CARBOHYD 816 816
SQ SEQUENCE 856 AA; 96938 MW; 0C241332CFE687 CRC64;

```

```

Query Match 75.7%; Score 112; DB 1; Length 856;
Best Local Similarity 85.7%; Pred. No. 3.3e-08;
Matches 24; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

```

OY 1 RVIRVQACRAIRHVRIRROGLRIL 28
Db 828 RVLEVQACRAIRHVRIRROGLRIL 855

```

```

RESULT 3
ENV_HVILW STANDARD; PRT; 861 AA.
AC P03377;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope glycoprotein gp160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].

```



```

FT CARBOHYD 289 289 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 295 295 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 301 301 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 332 332 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 339 339 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 356 356 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 386 386 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 392 392 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 397 397 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 406 406 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 448 448 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 463 463 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 611 611 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 616 616 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 624 624 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 637 637 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 674 674 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 750 750 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 816 816 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 856 AA; 97188 MM; 3373C68BB84C1AFC CRC64;

```

Query Match 69.6%; Score 103; DB 1; Length 856;
 Best Local Similarity 82.1%; Pred. No. 6e-07; Mismatches 0; Gaps 0;

```

Oy 1 RVIRVORACRAIRHIVRIRIGLRIL 28
    ||| ||| ||| ||| ||| ||| ||| |||
Db 828 RVIEVOEAVRAIRHIVRIRIGLRIL 855

```

```

RESULT 5
ENV_HV1ND STANDARD; PRT; 846 AA.
ID ENV_HV1ND
AC P18799:

```

```

DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (NDK isolate) (HIV-1).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11695;

```

```

RN SEQUENCE FROM N.A.
RX MEDLINE=90034200; Pubmed=2806917;
RA Spire B., Sire J., Zacher V., Rey F., Barre-Sinoussi F., Galibert F.,
  Hampe A., Chermann J.C.;
RT "Nucleotide sequence of HIV-1 NDK: a highly cytopathic strain of the
  human immunodeficiency virus.";
RL Gene 81:275-284(1989).
CC -1- MISCELLANEOUS: NDK, ISOLATED FROM A ZAIRIAN PATIENT AFFECTED WITH
  AIDS. AND IS A HIGHLY CYTOPATHOGENIC STRAIN.

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).

```

```

CC EMBL: M27323; AAA44873.1; -
DR PIR: J00066; VCLJND.
DR HIV: M27323; ENV$NDK.
DR InterPro: IPR000328; Env_GP41.
DR Pfam: PF00516; GP120; 1.
DR BLOD: PF00517; GP41; 1.
KW AIDS: Coat protein; polyprotein; glycoprotein; Transmembrane;
  Signal.
FT SIGNAL, 1 29

```

```

FT CHAIN 30 501 EXTERIOR MEMBRANE GLYCOPROTEIN.
FT CHAIN 502 846 TRANSMEMBRANE GLYCOPROTEIN.
FT DISULFID 53 73 BY SIMILARITY.
FT DISULFID 118 200 BY SIMILARITY.
FT DISULFID 125 191 BY SIMILARITY.
FT DISULFID 130 152 BY SIMILARITY.
FT DISULFID 213 242 BY SIMILARITY.
FT DISULFID 223 234 BY SIMILARITY.
FT DISULFID 291 328 BY SIMILARITY.
FT DISULFID 374 435 BY SIMILARITY.
FT DISULFID 381 408 BY SIMILARITY.
FT CARBOHYD 87 87 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 129 129 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 151 151 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 179 179 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 182 182 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 229 229 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 236 236 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 257 257 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 271 271 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 284 284 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 290 290 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 351 351 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 382 382 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 388 388 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 392 392 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 395 395 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 401 401 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 438 438 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 451 451 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 452 452 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 601 601 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 606 606 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 615 615 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 627 627 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 846 AA; 96476 MM; 8A3B9DA527DE2E83 CRC64;

```

Query Match 68.9%; Score 102; DB 1; Length 846;
 Best Local Similarity 75.0%; Pred. No. 8.2e-07; Mismatches 21; Conservative 3; Indels 4; Gaps 0;

```

Oy 1 RVIRVORACRAIRHIVRIRIGLRIL 28
    ||| ||| ||| ||| ||| ||| ||| |||
Db 818 RVIEVOEAVRAIRHIVRIRIGLRIL 845

```

```

RESULT 6
ENV_HV1W2 STANDARD; PRT; 847 AA.
ID ENV_HV1W2
AC P05880:

```

```

DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (WMJ2 isolate) (HIV-1).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11705;

```

```

RN SEQUENCE FROM N.A.
RX MEDLINE=86235450; Pubmed=3012778;
RA Hahn B.H., Shaw G.M., Taylor M.E., Redfield R.R., Markham P.D.,
  Salahuddin S.Z., Wong-Staal F., Gallo R.C., Parks E.S., Parks W.P.;
RT "Genetic variation in HIV-1/LAV over time in patients with AIDS or
  at risk for AIDS.";
RL Science 232:1548-1553(1986).
CC -1- MISCELLANEOUS: ISOLATES WMJ1, WMJ2, AND WMJ3 WERE OBTAINED FROM
  BLOOD SAMPLES SEQUENTIALLY TAKEN FROM A TWO-YEAR OLD HAITIAN WHO
  WAS PERINATALLY INFECTED BY HER MOTHER.

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

[illegible]

FT CARBOHYD 814 814 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SO SEQUENCE 853 AA; 96912 MW; 3377B93B6F22ABA CRC64;
 Query Match 68.2%; Score 101; DB 1; Length 853;
 Best Local Similarity 82.1%; Pred. No. 1.le-06;
 Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 Oy 1 RVIRVOCARAIHIVIRIIRIGLERIL 28
 ||| ||| | ||||| ||||| |||
 Db 826 RVIEVOCAYRAIRHIRIRIIRIGLERIL 853
 RESULT 8
 ENV_HY1B1 STANDARD; PRT; 856 AA.
 AC P03375; ID ENV_HY1B1
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Envelope glycoprotein GP160 precursor [contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
 GN ENV.
 OS Human immunodeficiency virus type 1 (BH10 isolate) (HIV-1).
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
 OX NCBI_Taxid=11678;
 RN [1]
 RP MEDLINE=8511123; PubMed=2578615;
 RA Ratner L., Haseltine W., Patarca R., Livak K.J., Starcich B.R.,
 RA Josephs S.F., Doran E.R., Ratselski J.A., Whitehorn E.A.,
 RA Baemster K., Ivanoff L., Pelteway S.R. Jr., Pearson M.L.,
 RA Latenberger J.A., Papas T.S., Grayeb J., Chang N.T., Gallo R.C.,
 RA Wong-Staal F.;
 RT "Complete nucleotide sequence of the AIDS virus, HTLV-III.";
 RL Nature 313:277-284(1985).
 RN [2]
 RP DISULFIDE BONDS, AND CARBOHYDRATE-LINKAGE SITES.
 RX MEDLINE=90285159; PubMed=2355006;
 RA Leonard C.K., Spellman M.W., Riddle L., Harris R.J., Thomas J.N.,
 RA Gregory T.J.;
 RT "Assignment of intrachain disulfide bonds and characterization of potential glycosylation sites of the type 1 recombinant human immunodeficiency virus envelope glycoprotein (gp120) expressed in Chinese hamster ovary cells.";
 RL J. Biol. Chem. 265:10373-10382(1990).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; M15654; AAA44205.1; -;
 DR PIR; A03973; VCLJH3.
 DR HIV; M15654; ENVSBI102.
 DR InterPro: IPR000328; Env GP41.
 DR InterPro: IPR000777; GP120.
 DR Pfam; PF00516; GP120; 1.
 DR Pfam; PF00517; GP41; 1.
 KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal.
 KW SIGNAL.
 FT CHAIN 1 30
 FT CHAIN 31 511
 FT CHAIN 512 856
 FT CHAIN 54 74
 FT DISULFID 119 205
 FT DISULFID 126 196
 FT DISULFID 131 157
 FT DISULFID 218 247
 FT DISULFID 228 239

FT DISULFID 296 331
 FT DISULFID 378 445
 FT DISULFID 385 418
 FT CARBOHYD 88 88
 FT CARBOHYD 136 136
 FT CARBOHYD 141 141
 FT CARBOHYD 156 156
 FT CARBOHYD 160 160
 FT CARBOHYD 186 186
 FT CARBOHYD 197 197
 FT CARBOHYD 230 230
 FT CARBOHYD 234 234
 FT CARBOHYD 241 241
 FT CARBOHYD 262 262
 FT CARBOHYD 276 276
 FT CARBOHYD 289 289
 FT CARBOHYD 295 295
 FT CARBOHYD 301 301
 FT CARBOHYD 332 332
 FT CARBOHYD 339 339
 FT CARBOHYD 356 356
 FT CARBOHYD 386 386
 FT CARBOHYD 392 392
 FT CARBOHYD 397 397
 FT CARBOHYD 406 406
 FT CARBOHYD 448 448
 FT CARBOHYD 463 463
 FT CARBOHYD 611 611
 FT CARBOHYD 616 616
 FT CARBOHYD 625 625
 FT CARBOHYD 637 637
 FT CARBOHYD 674 674
 FT CARBOHYD 750 750
 FT CARBOHYD 816 816
 SO SEQUENCE 856 AA; 97224 MW; 0BFEB1A18931BB27 CRC64;
 Query Match 68.2%; Score 101; DB 1; Length 856;
 Best Local Similarity 82.1%; Pred. No. 1.le-06;
 Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 Oy 1 RVIRVOCARAIHIVIRIIRIGLERIL 28
 ||| ||| | ||||| ||||| |||
 Db 828 RVIEVOCAYRAIRHIRIRIIRIGLERIL 855
 RESULT 9
 ENV_HY1PV STANDARD; PRT; 856 AA.
 ID ENV_HY1PV
 AC P03376; DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Envelope glycoprotein GP160 precursor [contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
 GN ENV.
 OS Human immunodeficiency virus type 1 (PV22 isolate) (HIV-1).
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
 OX NCBI_Taxid=11700;
 RN [1]
 RP MEDLINE=8511157; PubMed=2982104;
 RA Muesing M.A., Smith D.H., Cabradilla C.D., Benton C.V., Laszky L.A.,
 RA Capon D.J.;
 RT "Nucleic acid structure and expression of the human AIDS/lymphadenopathy retrovirus.";
 RL Nature 313:450-458(1985).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial

DB	828	RVEVYQRCRAIHHPRIIRIGLERAL	855	
RESULT	11			
ID	ENV_HV1B8	STANDARD:	PRT:	851 AA.
AC	P04582:			
DT	13-AUG-1987	(Rel. 05, Created)		
DT	13-AUG-1987	(Rel. 05, Last sequence update)		
DT	15-JUL-1999	(Rel. 38, Last annotation update)		
DE	Envelope polypeptide GP120 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].			
DE	ENV.			
OS	Human immunodeficiency virus type 1 (BH8 isolate) (HIV-1).			
OC	Viruses; Retroid viruses; Retroviridae; Lentiviruts.			
OX	NCBI_TaxID=11684;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=85111123; PubMed=2578615;			
RA	Ratner L., Haseltine W., Patarca R., Livak K.J., Starich B.R.,			
RA	Josephs S.F., Doran E.R., Palatka J.A., Whitehorn E.A.,			
RA	Baumenster K., Ivanoff L., Peteway S.R. Jr., Pearson M.L.,			
RA	Laubenberger J.A., Papas T.S., Chrayab J., Chang N.T., Gallo R.C.,			
RA	Wong-Straal F.;			
RL	"Complete nucleotide sequence of the AIDS virus, HTLV-III.";			
RL	Nature 313:277-284(1985).			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL: K02011; AAA44661.1; -.			
DR	HIV, K02011; ENVSEB8.			
DR	GlycoSuiteDB; P04582; -.			
DR	InterPro: IPR000328; Env_GP41.			
DR	InterPro: IPR000777; GP120.			
DR	Pfam: PF00516; GP120; 1.			
DR	Pfam: PF00517; GP41; 1.			
KW	AIDS; Coat protein; Polypeptide; Glycoprotein; Transmembrane;			
KW	Signal.			
FT	SIGNAL	1	30	
FT	CHAIN	31	506	EXTERIOR MEMBRANE GLYCOPROTEIN.
FT	CHAIN	507	851	TRANSMEMBRANE GLYCOPROTEIN.
FT	DISULFID	54	74	BY SIMILARITY.
FT	DISULFID	119	205	BY SIMILARITY.
FT	DISULFID	126	196	BY SIMILARITY.
FT	DISULFID	131	157	BY SIMILARITY.
FT	DISULFID	218	247	BY SIMILARITY.
FT	DISULFID	228	239	BY SIMILARITY.
FT	DISULFID	296	331	BY SIMILARITY.
FT	DISULFID	378	440	BY SIMILARITY.
FT	DISULFID	385	413	BY SIMILARITY.
FT	CARBOHYD	88		N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	136	136	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	141	141	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	156	156	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	160	160	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	186	186	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	197	197	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	230	230	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	234	234	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	241	241	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	262	262	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	276	276	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	295	295	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	301	301	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	332	332	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	339	339	N-LINKED (GLCNAC. . .) (POTENTIAL).

Query Match	Best Local Similarity	Score 99;	DB 1;	Length 851;
Matches 22;	Conservative 1;	Mismatches 5;	Indels 0;	Gaps 0;
QY 1 RIRIVYQRCRAIRIRIVRIRIOGLRLRL 28	111 : 111 111111 111111 111			
Db 823 RVIELVQAYRAIRIRIRIOGLRLRL 850				
RESULT 12				
ENV_HV122	STANDARD:	PRT:	853 AA.	
AC P12487:				
DT 01-OCT-1989 (Rel. 12, Created)				
DT 01-OCT-1989 (Rel. 12, Last sequence update)				
DT 16-OCT-2001 (Rel. 40, Last annotation update)				
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].				
GN ENV.				
OS Human immunodeficiency virus type 1 (Z2/CDC-Z34 isolate) (HIV-1).				
OC Viruses: Retroid viruses; Retroviridae; Lentivirus.				
OX NCBI_TaxID=11683;				
RN (1)				
RP SEQUENCE FROM N.A.				
RA Theodore T., Buckler-White A.;				
RL Submitted (NOV-1988) to the HIV data bank.				
CC -----				
CC This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC the European Bioinformatics Institute. There are no restrictions on its				
CC use by non-profit institutions as long as its content is in no way used				
CC modified and this statement is not removed. Usage by and for commercial				
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/				
CC or send an email to license@isb-sib.ch).				
CC -----				
DR EMBL; M22639; AAA45370.1; -.				
DR HIV; M22639; ENV5226.				
DR InterPro; IPR000328; ENV_GP41.				
DR InterPro; IPR000777; GP120.				
DR Pfam; PF00517; GP120; 1.				
DR AIDS; Coat protein; Glycoprotein; Transmembrane;				
KW Signal.				
FT SIGNAL 1 31				
FT CHAIN 32 508				
FT CHAIN 509 853				
FT DIISOLFD 53 73				
FT DIISOLFD 118 206				
FT DIISOLFD 125 197				
FT DIISOLFD 130 154				
FT DIISOLFD 219 248				
FT DIISOLFD 229 240				
FT DIISOLFD 297 330				
FT DIISOLFD 376 442				
FT DIISOLFD 383 415				
FT CARBOHYD 87 87				
FT CARBOHYD 137 137				
FT CARBOHYD 137 137				


```

FT CARBOHYD 144 144 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 153 153 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 157 157 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 185 185 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 188 188 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 198 198 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 235 235 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 242 242 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 263 263 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 277 277 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 290 290 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 296 296 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 331 331 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 338 338 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 353 353 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 384 384 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 390 390 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 402 402 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 441 441 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 445 445 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 458 458 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 459 459 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 462 462 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 608 608 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 613 613 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 622 622 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 634 634 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 671 671 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 853 AA; 97043 MW; 849B08BCBAFF7008 CRC64;

```

```

Query Match 66.2%; Score 98; DB 1; Length 853;
Best Local Similarity 67.9%; Pred. No. 3e-06;
Matches 19; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

```

```

OY 1 RVIRVORACRAIRHVRIRRIROGLRRL 28
Db 825 RVIEIIRACRAVLNIPRIRRIROGLRSL 852

```

```

RESULT 13
ENV_HV1EL STANDARD; PRT; 853 AA.
AC P04581;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (ELI isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_Taxid=11689;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86245056; PubMed=2424612;
RA Allison M., Main-Hobson S., Montagnier L., Sonigo P.;
RT "Genetic variability of the AIDS virus: nucleotide sequence analysis
of two isolates from African patients.";
RL Cell 46:63-74(1986).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; K03454; AAA44329.1; -
CC EMBL; A07108; CAA00616.1; -
CC HIV; K03454; ENV$ELI.
CC InterPro: IPR000328; Env_GP41.

```

```

DR InterPro: IPR000777; GP120.
DR Pfam: PF00516; GP120; 1.
DR Pfam: PF00517; GP41; 1.
KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane;
Signal.
FT SIGNAL 1 31
FT CHAIN 32 508
FT CHAIN 509 853
FT DISULFID 53 73
FT DISULFID 118 206
FT DISULFID 125 197
FT DISULFID 130 154
FT DISULFID 219 248
FT DISULFID 229 240
FT DISULFID 297 330
FT DISULFID 376 442
FT DISULFID 383 416
FT CARBOHYD 87 87
FT CARBOHYD 129 129
FT CARBOHYD 137 137
FT CARBOHYD 143 143
FT CARBOHYD 153 153
FT CARBOHYD 157 157
FT CARBOHYD 183 183
FT CARBOHYD 188 188
FT CARBOHYD 198 198
FT CARBOHYD 235 235
FT CARBOHYD 242 242
FT CARBOHYD 263 263
FT CARBOHYD 277 277
FT CARBOHYD 290 290
FT CARBOHYD 331 331
FT CARBOHYD 353 353
FT CARBOHYD 384 384
FT CARBOHYD 390 390
FT CARBOHYD 394 394
FT CARBOHYD 400 400
FT CARBOHYD 405 405
FT CARBOHYD 406 406
FT CARBOHYD 411 411
FT CARBOHYD 445 445
FT CARBOHYD 458 458
FT CARBOHYD 459 459
FT CARBOHYD 462 462
FT CARBOHYD 608 608
FT CARBOHYD 613 613
FT CARBOHYD 622 622
FT CARBOHYD 634 634
SQ SEQUENCE 853 AA; 96721 MW; F9CD864DAAD07A5 CRC64;

```

```

Query Match 64.9%; Score 96; DB 1; Length 853;
Best Local Similarity 67.9%; Pred. No. 5.7e-06;
Matches 19; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

```

```

OY 1 RVIRVORACRAIRHVRIRRIROGLRRL 28
Db 825 RVIEIIRACRAVLNIPRIRRIROGLRSL 852

```

```

RESULT 14
ENV_HV1A2 STANDARD; PRT; 855 AA.
AC P03378;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (ARV2/SF2 isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_Taxid=11685;

```


FT	CARBOHYD	276	276	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	289	289	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	295	295	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	301	301	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	331	331	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	354	354	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	360	360	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	384	384	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	390	390	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	396	396	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	400	400	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	442	442	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	456	456	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	607	607	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	612	612	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	621	621	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	633	633	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	670	670	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	812	812	N-LINKED (GLCNAC. . .) (POTENTIAL).
SO	SEQUENCE	852 AA;	97203 MM;	2BB86345DEC915F CRC64;

Query Match

Best Local Similarity 72.4%; Score 93; DB 1; Length 852;

Matches 21; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

OY 1 RVIRVVQACRAIRHIVRIQGLRLIR 29
DB 824 RAIEVVQAFRAIRHIVRIQGLRLIR 852

Search completed: August 14, 2002, 10:59:45
Job time: 501 sec

